

FIGURE 48

GCACGTTGGGCACCGCCATCAGCTGTTTCGCGCTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTGCG
TTGCATCTTTACACACTACAGCTATTGTTAGGTTGCTTCGGGACACAGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGGCTGGATCCTGTTT
TTCGTGCTCTATGATTCTGCAATTGTTTGATACACCCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCTCTGTG
GACAGCGTGGCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCTCACTGCTGCCACTGCTTTGAAAAGGCAGCAGCAACAGAATCTGAATTCCTGGT
CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGGCCAGGGCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCTGCCCCAGGCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAATCTGATCTACAACACAGT
GCACCAGCGACACCTGTCAAACCCGGCCCGGCTGGGATGCTATGTGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTGAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTCTGGCTGCAGGCTCAGATTGAGGGGGCAGCTT
TCCTGGCCACAGAGCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCTCCAGGCAGGAGCACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGAGACGTAGGGCTGGGACCCAGA
CCGAGGAGTGGGGCTGAAGCAGCTCATCTGCATGGAGCCTACACCCACCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCTGCCCTATCCTGACCACACCTGCCTGATGGGAGCGTGGCTGGGTTCTGGGACCG
GCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCCTCCTGGGGCCTAG
GGCTGCAGCCGGTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACAGTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG
CATGAGGTGAGGGGCATATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTCAACCGCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAAATGGCATTACTGCCCTCTGCTCCCAACCTGTCTATGTGTGATTCCAGGCAC
CAGGGCAGGCC CAGAAGCCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCAACCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAATCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCTTACTCTTTCAGATACAATCAGCCAGCCACGTTGTTT
TGAAAATTTCTTTTGGGGGGCAGCAGTTTTCCTTTTAACTTAAATAAATGTTTAC
AAAAATAAA

09978295.101501

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQDSSGGFVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAPLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMAILLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVPVVTLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSPGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

007825.101504

THE UNIVERSITY OF CHICAGO

CGGGCCGCCCCCGGCCCCCATTCGGGCGGGGCTCGTGTGGCGGGGACTGAGCCAGGCTTG
GCCGCGTCCCTGAGTCCCAGAGTCCGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTGACGCCCTCACAGGAAGATGCTGCGTGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCTCTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAACA
GCTGGTGCACAGCTTTGTGTAGGGGCCAGGACCAGGGCAGCGCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCTGCTGTGGCG
GACGAGGGGAGCTTCACTGCTTCGTGAGCATCCGGGATTTCCGCAGCGCTGCCGTGAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCAGCATGACCCTGGAGGCCAACAAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCTTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCTTGACTGGCAACGTGACCAGCTCGCAGATGGCCAACGA
GACGGGCTGTGTTTGTATGTCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCTCTGTGTGCGCAACCCCGTGCAGTGCAGGATGCGCACCRGCTGTGTACCATCACAGGG
CAGCCCTATGATCATCCCCCTGAGGCGCCCTGTGGTGACCGTGGGGCTGTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCGTGTCTGGAGAAGATCAAAACAGCCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGAACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCTACCCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC
AGACCACTGTGACAGCCTTATTCTCCAATGGACATGATTCCAAGTCATCTGTGCTGCTTTT
TTCTTATAGACACAATGAACAGACCACCACAACCTTAGTCTCTAAGTCATCTGCCTGCT
GCCTTATTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCTC
TTCTTCCAGTGTGCGTGGACCATCTGGCTGCCCTTTTTTCTCCAAAGATGCAATATTGAGA
CTGACTGACCCCCCTGCCCTATTTCACCAAAGACACGATGCATAGTCAACCCGGCCTTGTTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTGAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCCTTC
CCCTCCTTCTCTCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCCACCCCCACATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCCTGGC
TTGCCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCCTCCG
GATGTCACTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAA